THE BALTIC JOURNAL
OF ROAD AND BRIDGE ENGINEERING
2025/20(3)

ISSN 1822-427X/eISSN 1822-4288
2025 Volume 20 Issue 3: 57–89
https://doi.org/10.7250/bjrbe.2025-20.662

THE EVOLUTION OF CLASSICAL AND SOFT COMPUTING METHODS IN PREDICTING ROAD MAINTENANCE AND REPAIR COSTS: APPROACHES IN THE LITERATURE AND FUTURE PERSPECTIVES

HAYDAR GUNDOGDU*, OMER FARUK CANSIZ, MEHMET FATIH CAN

Graduate Education Institute, Iskenderun Technical University, Iskenderun, Turkey

Received 16 January 2025; accepted 2 September 2025

Abstract. Road infrastructure is critical to the economic, social, and environmental sustainability of modern societies. This study compares classical methods (Multiple Linear Regression, Ridge, and LASSO) with soft computing techniques (Artificial Neural Networks, Fuzzy Logic, Random Forests, Gradient Boosting, Support Vector Machines, and Genetic Algorithms) for predicting road maintenance and repair costs. A comprehensive search has been conducted in Web of Science, and Scopus for studies published between January 2010 and March 2024. Boolean operators and specific key terms such as "road maintenance costs," "soft computing," and "classical prediction methods" have been used. The approach has been PRISMA-inspired but adapted for narrative review purposes; hence, no formal quality assessment or meta-analysis has been performed. Peer-reviewed journal articles have been included, while grey literature has been excluded to ensure methodological consistency. While classical methods offer simplicity and computational efficiency, they often fall short in addressing complex data structures such as non-linear relationships and multicollinearity. Conversely, soft computing techniques excel in modelling non-linear systems and managing uncertainties. Hybrid models combining classical and soft computing approaches enhance

* Corresponding author. E-mail: hdmgundogdu@gmail.com

Haydar GUNDOGDU (ORCID ID 0000-0002-2354-8920)
Omer Faruk CANSIZ (ORCID ID 0000-0001-6857-2513)
Mehmet Fatih CAN (ORCID ID 0000-0002-3866-2419)

prediction accuracy by 20–30%, providing improved capabilities in modelling environmental factors. However, further research is required to evaluate their long-term performance and adaptability to diverse geographical conditions. This study highlights the theoretical advantages of hybrid models while offering practical solutions for sustainable infrastructure management. The findings provide policymakers and engineers with actionable insights, promoting efficient public resource use and sustainable development goals. Future research should focus on integrating IoT and big data analytics to address dynamic environmental variables, fostering innovation in infrastructure management.

Keywords: artificial neural networks, hybrid models, LASSO, narrative review, ridge, repair costs, road maintenance, soft computing techniques, sustainable infrastructure management.

Introduction

The maintenance and repair of road infrastructure are critical to the economic, social, and environmental sustainability of modern societies. Roads serve as essential connectors for industries such as logistics, trade, and tourism, as well as for daily commuting, underscoring their indispensable role in supporting economic growth and societal well-being. As populations increase, urbanization accelerates, and global trade volumes expand, the demand for more resilient and efficiently managed road networks continues to grow. Within this context, accurate estimation of road maintenance and repair costs emerges as a fundamental requirement for sustainable infrastructure management and effective allocation of public resources.

A considerable body of literature has explored methods for predicting road maintenance and repair costs. Classical approaches, including Multiple Linear Regression (MLR), Ridge Regression, and LASSO, have long been utilised for their simplicity, computational efficiency, and accessibility. However, these methods are constrained by their reliance on linear assumptions and fixed parameters, making them less effective in handling non-linear relationships and complex datasets characterised by multicollinearity. Multicollinearity, which arises from high correlations among independent variables, is a particularly significant challenge, as it reduces model reliability and predictive accuracy.

In contrast, soft computing techniques have garnered increasing attention for their ability to model complex, non-linear systems and manage uncertainties. Methods such as Artificial Neural Networks (ANNs), Random Forests (RFs), Gradient Boosting, Fuzzy Logic, and Genetic Algorithms (GAs) offer greater flexibility and adaptability compared to classical approaches. These techniques have demonstrated their effectiveness in addressing the limitations of traditional methods, particularly in scenarios involving intricate interdependencies among variables and the integration of environmental and socio-economic factors into predictive models.

Hybrid models that combine classical and soft computing techniques represent an emerging trend in the literature, addressing the limitations of individual approaches while leveraging their strengths. For example, the capacity of Ridge Regression to handle multicollinearity can be combined with the non-linear modelling capabilities of Artificial Neural Networks, yielding more robust and accurate predictions. Studies indicate that hybrid approaches can enhance predictive accuracy by 20–30%, offering significant improvements in the context of infrastructure management. These models also facilitate the inclusion of dynamic factors such as traffic density, climate variability, and material quality, which are often overlooked in conventional frameworks.

In countries with extensive and complex road networks, such as Turkey, road infrastructure management is a key priority. The General Directorate of Highways (KGM) provides comprehensive datasets on personnel costs, material expenditures, energy consumption, and asphalt surface repairs, offering unique opportunities for developing advanced predictive models. The KGM's own maintenance strategy framework, which categorises interventions into five distinct levels based on road condition – from "no maintenance needed" to "road reconstruction required" (Sağlık & Güngör, 2008) – provides a valuable structured taxonomy that can be used to inform and validate the output of predictive models. These datasets allow for detailed analysis of road maintenance and repair costs, highlighting the importance of integrating both classical and soft computing techniques to address the unique challenges posed by dynamic economic and environmental conditions.

This study makes an original contribution by providing a systematic, side-by-side evaluation of classical, soft computing, and hybrid approaches for road maintenance cost prediction – an integration rarely addressed collectively in prior research. Beyond synthesising methodological strengths and limitations, it uniquely incorporates insights from Turkey's extensive road network data while identifying forward-looking strategies, including the integration of IoT and big data analytics, to enhance adaptability under dynamic environmental and socio-economic conditions. By bridging theoretical rigor with practical applicability, this work offers a novel framework for developing more accurate, sustainable, and context-sensitive infrastructure management solutions.

This study aims to conduct a comprehensive literature review on classical and soft computing methods used for predicting road maintenance and repair costs. By critically examining the strengths, limitations, and applications of these methods, the research seeks to establish a theoretical foundation and identify opportunities for methodological innovation. The inclusion of hybrid approaches as a focal point provides a pathway for addressing existing gaps in the literature while offering actionable insights for practitioners and policymakers.

The tasks of this literature review are:

– To assess the performance of classical methods, such as Multiple Linear Regression, Ridge, and LASSO, in terms of computational efficiency and interpretability;

– To evaluate the accuracy and versatility of soft computing techniques, including Artificial Neural Networks, RFs, and Gradient Boosting, in managing non-linear and complex datasets;
– To explore the potential of hybrid models in integrating classical and soft computing methods for improved predictive performance;
– To identify gaps in the existing literature and propose future research areas that address emerging challenges in infrastructure management;
– To provide a comprehensive theoretical and practical framework for decision-makers in the transportation sector, fostering sustainable infrastructure development.

This study contributes to the growing body of research on infrastructure management by presenting a comparative analysis of classical and soft computing methods, highlighting their respective strengths and weaknesses. By focusing on hybrid models, the research not only bridges the gap between traditional and contemporary methodologies but also offers a robust foundation for sustainable infrastructure management. The findings have far-reaching implications, providing a roadmap for policymakers, engineers, and researchers to develop innovative solutions for the efficient and resilient management of road networks.

In conclusion, this literature review establishes a comprehensive perspective on the evolution of predictive methods for road maintenance and repair costs, setting the stage for future advancements in the field. The integration of classical and soft computing techniques within hybrid models has the potential to revolutionise infrastructure management practices, ensuring long-term sustainability and efficiency in transportation systems.

1. Methods

As part of this study, an extensive literature review has been conducted using multiple high-impact databases. The databases utilised include internationally recognised academic sources such as Web of Science and Scopus between January 2010 and March 2024. These platforms have been selected for their ability to provide up-to-date and reliable research on road infrastructure management and maintenance costs.

Key terms have been carefully selected to review relevant literature. Keywords such as "road maintenance costs," "soft computing," "classical prediction methods," "infrastructure management," and "artificial neural networks" have been employed. Boolean operators (e.g., "AND," "OR") were used to combine keywords and generate more specific search results. For instance, phrases like "road AND soft computing" have been applied as part of a comprehensive search strategy to identify both specialised and related studies. Example search query in Scopus:

TITLE-ABS-KEY("road maintenance costs" OR "road repair costs") AND TITLE-ABS-KEY("soft computing" OR "artificial neural networks" OR "machine learning") AND TITLE-ABS-KEY("classical methods" OR "multiple linear regression" OR "ridge regression" OR "lasso regression") and in Web of Science: TS=("road maintenance costs" OR "road repair costs") AND TS=("soft computing" OR "artificial neural networks" OR "machine learning") AND TS=("classical methods" OR "multiple linear regression" OR "ridge regression" OR "lasso regression"). The flowchart of the searching process is illustrated in Figure 1.

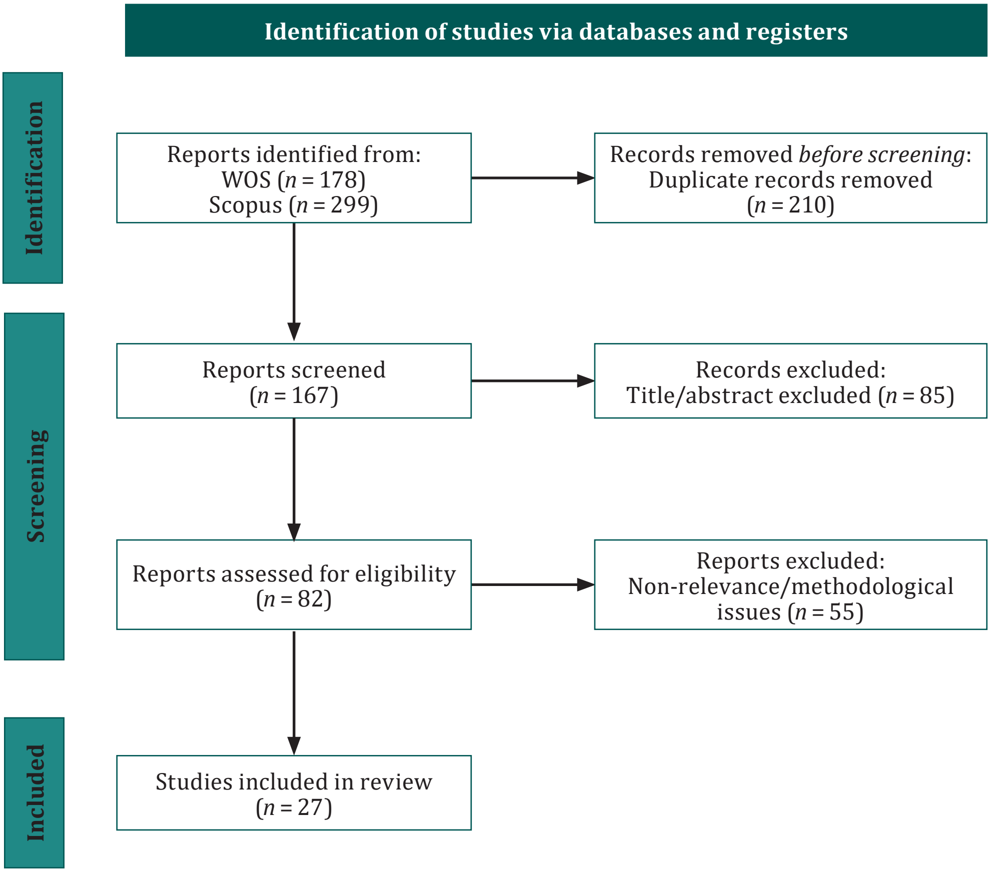

Figure 1. Flowchart of study selection

This study has been conducted following a process designed to align with academic standards for literature reviews. The scope includes peer-reviewed publications available in full text, covering methods for predicting road maintenance and repair costs. Detailed descriptions of the methods used in the studies, their accuracy metrics, and comparative analysis results have been examined. Publications in both English and Turkish have been included in the review without language restrictions.

Inclusion criteria:

– Studies employing classical methods (e.g., Multiple Linear Regression, Ridge, LASSO) and soft computing techniques (e.g., Artificial Neural Networks, RFs, Gradient Boosting, Fuzzy Logic, Genetic Algorithms) for predicting road maintenance and repair costs;
– Research that thoroughly addresses the accuracy, advantages, and limitations of these methods;
– Full-text, peer-reviewed articles published in journals or presented at high-impact conferences.

Exclusion criteria:

– Publications in abstract or short-paper formats;
– Grey literature (theses, reports);
– Studies focusing on non-road infrastructure types or other industries;
– Publications lacking adequate methodological explanation or containing unverified findings.

These criteria have been established to ensure a robust foundation for the reliability and focus of the literature review. While no specific database search protocol has been strictly adhered to, the process has been structured around an academic framework. This review follows a PRISMA-inspired (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) approach tailored for a narrative review design. This methodology has been adopted to enhance methodological consistency and ensure repeatability for future researchers.

1.1. Factors influencing road maintenance and repair costs

Analysing the factors affecting road maintenance and repair costs is crucial for developing infrastructure management strategies and ensuring sustainable transportation systems. Elements such as traffic density, heavy vehicle loads, environmental conditions, material quality, and the ageing of infrastructure are key determinants of these costs. While many studies in the literature focus on each of these variables individually, a deeper examination of their interactions is necessary.

1.1.1. Impact of traffic density and heavy vehicle loads

Heavy vehicles exert a significant impact on road maintenance costs and pavement deterioration. Gibby et al. (1990) observed that a single heavy truck caused damage equivalent to that of 90 passenger cars, with annual maintenance costs amounting to $7.60 per mile compared to $0.08 per mile for passenger vehicles. Similarly, Costanzi and Cebon (2007) demonstrated that the adoption of road-friendly suspension systems had the potential to reduce maintenance expenditures by 14%. However, they also noted that poorly maintained shock absorbers could exacerbate these costs by as much as 21%, highlighting the critical importance of vehicle upkeep in mitigating road damage. Pais et al. (2019) further emphasised the detrimental effects of overloaded vehicles, revealing that such practices increased pavement damage and elevated life cycle costs by approximately 30%. Ahmad (2023) underscored the value of proactive maintenance programs, coupled with data-driven decision-making, in optimising the longevity and cost-efficiency of highway infrastructure. By enabling the timely identification and resolution of potential issues, these strategies can significantly reduce the long-term costs associated with road maintenance. Collectively these studies underscore the substantial impact of heavy vehicles on road infrastructure. They highlight the necessity of implementing effective maintenance strategies, leveraging innovative technologies, and enacting policy measures to ensure the efficient management of road assets. Such approaches not only prolong the lifespan of road networks but also contribute to broader economic and environmental sustainability.

1.1.2. Role of environmental factors

Environmental factors have a profound impact on the sustainability and performance of road infrastructure. Climate change and air pollution play a critical role in the degradation of construction materials; temperature fluctuations, freeze-thaw cycles, and rainfall cause detrimental effects on road surfaces and substructures (Kumar & Imam, 2013). To mitigate these effects, the implementation of effective drainage systems is essential. Proper drainage infrastructure can significantly reduce the adverse effects of water-related issues, such as flooding and heavy rainfall, which often result in cracks, structural failures, and eventual collapses in road systems (Awwad, 2021).

Life Cycle Assessment (LCA) studies focusing on pavement infrastructure highlight the necessity of developing more comprehensive and consistent methodologies. Accurate comparisons of the environmental impacts of different pavement types require consideration of performance-adjusted functional units, region-specific data, and standardised impact categories (Inyim et al., 2016). These factors are crucial to ensure that LCA outcomes accurately reflect the long-term environmental and functional performance of various pavement designs.

Furthermore, the impact of climate on geotechnical infrastructure, such as transport networks and flood defense systems, is significant. Climate-induced failures in these systems can lead to major disruptions and economic losses, underscoring the necessity for resilient design and construction practices (Vardon, 2015). Addressing these challenges demands the integration of climate-resilient strategies into infrastructure planning and development to enhance the durability and sustainability of road networks and associated geotechnical systems.

1.1.3. Material quality and service life

Recent research underscores the critical importance of material quality and sustainability in road construction practices. The incorporation of recycled materials into asphalt mixtures presents substantial benefits, including enhanced cost efficiency, environmental sustainability, and improved overall performance (Bamigboye et al., 2021). Evidence suggests that recycled asphalt pavements (RAP) can exhibit superior mechanical properties compared to conventional mixtures. For instance, Sapkota et al. (2023) reported that recycled asphalt mixtures demonstrated up to a 145% increase in stiffness and a 99% higher Marshall stability, indicative of their enhanced structural integrity. Moreover, these mixtures exhibit improved resistance to water-induced damage, thereby significantly enhancing their durability and lifespan (Sapkota et al., 2023). Although the upfront costs associated with concrete pavements may be higher, they offer notable long-term advantages in terms of durability and reduced maintenance requirements, particularly for high-traffic highways. Concrete pavements have been recognised for their resilience under heavy loads, making them an optimal choice for infrastructure subjected to intense and continuous vehicular stress (Soares Santos et al., 2024). The integration of recycled materials into road construction not only addresses critical waste management challenges but also alleviates the dependency on natural aggregates. This shift leads to notable reductions in life cycle environmental impacts and associated costs, fostering sustainable construction practices (Bamigboye et al., 2020; AlKheder, 2022). By reducing the demand for virgin materials and promoting circular economy principles, these approaches contribute to a more environmentally responsible and economically viable construction paradigm, aligning with global efforts to enhance infrastructure sustainability.

1.1.4. Ageing infrastructure and the need for periodic maintenance

Recent literature highlights the growing importance of green infrastructure (GI) and maintenance in sustainable urban development. GI has been identified as a key strategy for integrating environmental, social, and economic growth, particularly in addressing the impacts of climate change (Ying et al., 2022; Fassa, 2024). Studies emphasise the need for regular maintenance of ageing infrastructure, with older roads incurring significantly higher costs (Nyame, 2024). Life Cycle Assessments of road and rail infrastructure increasingly consider maintenance, with various approaches used to estimate maintenance frequency and account for the effects of climate change (Liljenström et al., 2022). Research on GI has focused on stormwater management, ecosystem services, biodiversity protection, and climate change adaptation (Ying et al., 2022). Key areas requiring further investigation include green materials, water and flood management solutions, urban heat mitigation, and GI regulations, particularly in developing countries such as Indonesia (Fassa, 2024).

1.2. Applications of classical estimation methods in road maintenance and repair costs

Classical estimation methods have long been fundamental tools for calculating road maintenance and repair costs. Methods such as multiple linear regression (MLR), Ridge regression, and LASSO are widely utilised in such analyses. These approaches offer advantages in terms of accuracy, computational speed, and model simplicity. However, they may fall short in dealing with complex datasets and systems with non-linear relationships. This section examines the applications of classical methods in road maintenance and repair costs and evaluates findings from the literature.

1.2.1. Use of multiple linear regression and variants

Multiple linear regression is a widely utilised statistical method for modeling relationships between variables; however, it exhibits limitations when applied to high-dimensional data and non-linear relationships (Hu et al., 2019). To address these challenges, researchers have proposed various enhancements and alternatives. Li et al. (2023) incorporated Principal Component Analysis (PCA) with MLR to reduce dimensionality and mitigate multicollinearity, thereby improving prediction accuracy in applications such as stock price forecasting. Etemadi and Khashei (2021) introduced a novel MLR approach that prioritises model reliability over accuracy, achieving improved generalisation ability in 63.333% of cases across 30 benchmark datasets. These studies underscore the ongoing advancements in enhancing MLR's effectiveness for complex datasets and its applicability across diverse fields.

1.2.2. LASSO and Ridge regression methods

LASSO and Ridge regression are regularization methods used to address multicollinearity in regression analysis. LASSO, which uses L1 regularization, excels in variable selection by shrinking some coefficients to zero, enhancing model interpretability and prediction accuracy (Xi et al., 2023; Omer, 2022). Ridge regression, employing L2 regularization, introduces a small bias to improve long-term predictions and model stability (Omer, 2022). Both methods outperform traditional least squares regression in the presence of multicollinearity (Omer, 2022; Rahmawati & Suratman, 2022). Comparative studies have shown that LASSO generally performs well across different sample sizes and multicollinearity levels, while Elastic Net, a hybrid approach, may be more accurate with larger samples (Oyeyemi et al., 2015). Ridge regression tends to yield lower mean squared errors, indicating better overall performance, but the ability of LASSO to produce sparse models by eliminating irrelevant variables makes it superior for model interpretation in some cases (Rahmawati & Suratman, 2022).

1.2.3. Accuracy, computational speed, and cost advantages of classical methods

Recent studies have compared classical and machine learning methods for forecasting and regression tasks. Makridakis et al. (2018) concluded that traditional statistical methods outperformed machine learning approaches in terms of accuracy and computational efficiency across various forecasting horizons. Similarly, Reznychenko et al. (2024) reported that RFs generally achieved superior accuracy compared to neural networks, whereas logistic regression exhibited limitations when applied to complex datasets. Portnoy and Koenker (1997) demonstrated that L1 methods could be computationally competitive with L2 methods when working with large datasets. In fraud detection applications, NarayananM et al. (2023) emphasised the efficiency of the Alternating Direction Method of Multipliers (ADMM) in optimising LASSO and Ridge regression, significantly reducing computational time compared to traditional implementations. These findings indicate that classical methods frequently provide advantages in accuracy and speed for simpler datasets, while advanced techniques such as ADMM enhance the performance of regularized regression methods in addressing more complex problems.

1.2.4. Performance of classical methods on existing infrastructure datasets

Classical methods for cost estimation in infrastructure projects have demonstrated varying levels of performance across different datasets. Ridge regression and LASSO have proven effective in handling large datasets with low noise levels and selecting significant variables, respectively (Aloraini, 2017). LASSO outperformed ordinary least squares regression in forecasting highway construction costs, offering advantages such as automatic feature selection and the ability to handle correlated data (Zhang et al., 2017). A study on road construction in Iraq identified Ridge regression as the best-performing model, achieving near-perfect accuracy metrics (Abed et al., 2022). However, the performance of these methods is heavily influenced by dataset characteristics and prevailing economic factors (Aloraini, 2017; Zhang et al., 2017). While classical statistical techniques such as regression analysis and Monte Carlo simulation remain widely used, more advanced methods like neural networks and support vector machines (SVMs) have demonstrated superior performance in cost estimation models (Atapattu et al., 2022). Furthermore, hybrid models that combine multiple techniques have shown potential to enhance estimation accuracy and reliability. Soft computing techniques hold a significant place in predicting road maintenance and repair costs due to their innovative approaches in modelling non-linear relationships, managing uncertainties, and understanding complex systems. Methods such as ANNs, Fuzzy Logic, Genetic Algorithms, RFs, Gradient Boosting, and SVMs are employed to predict these costs accurately and effectively. Each soft computing method is detailed below and supported by prominent studies in the literature.

1.3. Applications of soft computing estimation methods in road maintenance and repair costs

1.3.1. Artificial Neural Networks

ANNs have demonstrated significant potential in predicting maintenance costs for various infrastructure projects, particularly in the transportation sector. Studies have highlighted the capacity of ANNs to accurately estimate annual maintenance expenditures for highways (Woldemariam et al., 2016) and predict structural costs for highway projects (Al-Zwainy & Aidan, 2017). ANNs have also been effectively applied to forecast performance budgets for road defects (Cabrera & Silva, 2023) and estimate repair costs for bridges (Bouabaz & Hamami, 2008). Comparative analyses indicate that Convolutional Neural Networks (CNNs) outperform traditional fully connected ANNs in predicting bridge maintenance costs (Wang et al., 2022). Moreover, ANNs have been utilised to develop automated decision-making systems for highway pavement preventive maintenance (Li et al., 2022) and to forecast tractor repair and maintenance costs (Rohani et al., 2011). These studies underscore the effectiveness of ANNs in modelling complex relationships and enhancing prediction accuracy in the context of infrastructure maintenance cost estimation.

1.3.2. Fuzzy logic

Fuzzy logic has emerged as a powerful tool for modelling uncertainty across various domains, including cost estimation, traffic management, and logistics optimisation. In highway construction, fuzzy logic models have been utilised to predict cost overruns and support decision-making processes (Thorat & Birajdar, 2015). Similarly, in software development, fuzzy logic approaches have enhanced the accuracy of cost and time estimation (Ravishankar & Latha, 2012). Traffic signal control systems have benefited from the application of fuzzy logic controllers, which have significantly improved junction management efficiency (Koukol et al., 2015). In the field of logistics, fuzzy logic has been employed to optimise transportation routing by accounting for uncertainties in variables such as traffic conditions and delivery times (Ferreira, 2019). Additionally, GAs have been integrated with fuzzy logic to further enhance prediction accuracy in applications including commodity pricing (Chen & Wang, 2017), software cost estimation (Hamdy, 2014), and mandatory lane-changing models (Hou et al., 2012). These studies underscore the versatility and efficacy of fuzzy logic in addressing complex, real-world challenges across a wide range of industries.

1.3.3. Genetic algorithms

GAs are robust tools for solving optimisation problems and optimising variable selection. Goldberg (2013) noted their effectiveness in producing results for complex datasets and improving prediction accuracy, especially in multi-variable systems like road maintenance costs. Smith et al. (2021) reported that GAs improved accuracy by 18% by optimising variable selection for road maintenance costs. GAs are particularly effective for datasets with non-linear and complex relationships. Genetic algorithms based on evolutionary mechanisms in nature are used to achieve successful results in complex optimization problems (Grefenstette, 2014, p. 23). A new hybrid method that adaptively integrates local search techniques into a genetic algorithm in road maintenance and repair strategy selection offers more effective results in the search process compared to the traditional genetic algorithm (Santos et al.,2019). Chen and Wang (2017) demonstrated that integrating GAs with other soft computing techniques (e.g., ANNs and Fuzzy Logic) reduced error rates by 12%. They noted significant performance improvements in hybrid models when GAs were used in high-dimensional datasets. Rehman et al. (2020) showed that combining GAs with backpropagation algorithms enhanced prediction accuracy in large-scale forecasting problems. In conclusion, GAs are a powerful tool for predicting complex systems like road maintenance costs. When combined with hybrid approaches, GAs provide superior performance in improving prediction accuracy and reducing error rates.

1.3.4. Random Forests

RF is a versatile machine learning algorithm extensively utilised for classification and regression tasks. By combining multiple decision trees, it enhances prediction accuracy and effectively handles complex datasets (Biau & Scornet, 2016). RF offers several advantages, including its ability to manage unbalanced data, accommodate missing values, and mitigate overfitting (Salman et al., 2024). The algorithm has been successfully applied across a range of domains, including road surface assessment (Raman et al., 2024), accident severity prediction (Adefabi et al., 2023), and road sign maintenance (Radhika et al., 2023). The effectiveness of RF in variable selection has been highlighted by Hapfelmeier and Ulm (2013) as well as Sandri and Zuccolotto (2006), underscoring its utility in identifying critical predictors. Furthermore, RF has demonstrated promise in road health monitoring, where it has been employed to classify different types of road damage and estimate repair costs (Dugalam & Prakash, 2024). The algorithm's versatility and robust performance make it a widely preferred choice for diverse machine learning applications.

1.3.5. Gradient Boosting

Gradient Boosting (GB) has emerged as a highly effective machine learning technique with diverse applications in transportation and infrastructure management. It has demonstrated superior performance in predicting highway-rail grade crossing crashes (Pan Lu et al., 2020), forecasting short-term traffic flow (Li et al., 2019), and estimating traffic speed classes (Menguc et al., 2023). GB has also been successfully utilised to analyse rural poverty incidence related to road infrastructure (Chan et al., 2022) and to study airport pavement deterioration (Barua et al., 2021). Moreover, the method has proven valuable in assessing the carbon footprint of urban road networks (Yu et al., 2024). Originally introduced by Friedman (2001) as a robust and interpretable approach for function approximation, GB has since been applied to a wide range of domains, including auto insurance loss cost modelling (Guelman, 2012). These studies underscore the capacity of GB to handle complex, non-linear relationships and its effectiveness in addressing various prediction tasks related to transportation and infrastructure management.

1.3.6. Support Vector Machines

SVMs have demonstrated significant success in predicting road maintenance costs and classifying pavement defects. Studies have highlighted the effectiveness of SMSs in estimating project costs with high accuracy (Nabil et al., 2014) and predicting pavement roughness (Georgiou et al., 2018). Hybrid models that combine SVMs with other techniques, such as the Bromilow time-cost model, have achieved even greater accuracy in cost prediction (Petrusheva et al., 2019). SVMs-based approaches have also been applied to traffic flow management (Naveed et al., 2024) and bridge evaluation (Li & Burgueño, 2010). The integration of SVMs with convolutional neural networks has further improved road defect classification (Soni & Gujar, 2024), while conditional SVMs have enhanced traffic management in smart cities (Manogaran et al., 2022). These soft computing methods have shown considerable promise in various aspects of construction sustainability, including cost prediction, energy consumption, and environmental monitoring (Petruseva et al., 2019), making them invaluable tools for road infrastructure management.

1.4. The multicollinearity problem and its impact on prediction models

Multicollinearity is a significant issue caused by strong correlations between independent variables, negatively affecting the performance of prediction models. This problem, frequently encountered in estimating road maintenance and repair costs, reduces prediction accuracy and limits model generalisability. This section explores the definition of the multicollinearity problem, solutions using classical and soft computing methods, variable selection strategies, and the impact on model performance.

1.4.1. Definition of multicollinearity and its role in road prediction models

Multicollinearity, defined as the high correlation among predictor variables in regression models, can significantly affect the reliability and interpretation of results (Alin, 2010; Daoud, 2017). While some studies argue that multicollinearity does not impact prediction accuracy (Morris & Lieberman, 2018), others highlight its adverse effects on coefficient stability and model interpretability (Paul, 2006; Owoyemi & Bolakale, 2024). Common detection methods include examining correlation matrices, variance inflation factors (VIFs), and eigenvalues (Schreiber-Gregory, 2018; Upendra et al., 2023). The consequences of multicollinearity include inflated standard errors and unreliable coefficient estimates (Daoud, 2017; Barakat, 2023). To address this issue, researchers recommend remedies such as ridge regression, principal component regression, and variable selection techniques (Paul, 2008; Schreiber-Gregory, 2018; Upendra et al., 2023). Understanding and mitigating multicollinearity is essential for developing robust regression models, particularly in fields such as road maintenance prediction, where predictor variables are often intrinsically interrelated (Owoyemi & Bolakale, 2024).

1.4.2. Addressing multicollinearity with classical methods

Multicollinearity in regression analysis can result in unstable coefficient estimates and unreliable outcomes. Classical methods such as Ridge regression and LASSO have proven effective in addressing this issue (Efeizomor, 2023; Owoyemi & Bolakale, 2024). Ridge regression performs exceptionally well in datasets with high correlation among variables, delivering more stable results and improved prediction accuracy (Abdulhafedh, 2022; Rahmawati & Suratman, 2022). LASSO, by contrast, is more suitable for datasets with numerous variables, as it can shrink some coefficients to zero, thereby enhancing model interpretability (Rahmawati

& Suratman, 2022; Schreiber-Gregory et al., 2018). The Elastic Net method, which integrates features of both Ridge and LASSO, has been shown to outperform either approach in scenarios characterised by varying degrees of multicollinearity (Altelbany, 2021). However, the choice of method should be guided by the specific structure of the dataset and sample size (Oyeyemi et al., 2015). These regularization techniques provide valuable alternatives to traditional ordinary least squares (OLS) regression when addressing the challenges posed by multicollinearity (Barakat, 2023).

1.4.3. Solving multicollinearity with soft computing methods

Multicollinearity in regression analysis presents significant challenges for accurate modelling and prediction. While traditional statistical techniques such as Ridge regression and Principal Component Analysis effectively address this issue (Mahmood, 2024; Garg & Tai, 2013), machine learning approaches offer promising alternatives. ANNs have demonstrated superior performance in handling multicollinearity compared to Ordinary Least Squares (OLS) regression (Obite et al., 2020). Similarly, RFs and Gradient Boosting algorithms have been shown to effectively mitigate the effects of multicollinearity (Chan et al., 2022). Nonetheless, some studies indicate that statistical models, particularly Liu regression, outperform machine learning algorithms in datasets with high levels of multicollinearity (Yıldırım, 2024). Ensemble methods involving neural networks can address multicollinearity through incremental, negatively correlated, optimal convex blending (Lagari et al., 2021). The selection of an appropriate method depends on the characteristics of the dataset and the research context. Ongoing research explores the integration of Explainable AI techniques to better understand and address multicollinearity issues (Salih, 2024).

1.4.4. Variable selection strategies to reduce multicollinearity

Variable selection strategies are essential for addressing multicollinearity in regression models. LASSO and its extensions have demonstrated significant advantages in variable selection, processing high-dimensional data, and resolving multicollinearity issues (Xi et al., 2023). Ridge regression and LASSO consistently outperform traditional methods such as Ordinary Least Squares (OLS) in handling multicollinearity and enhancing predictive accuracy (Efeizomor, 2023). GAs have also proven effective for variable selection, particularly in datasets with high multicollinearity (Kilinc et al., 2016). Comparative studies indicate that GAs achieve the lowest error rates, while LASSO is most effective in reducing the number of variables (Ryu et al., 2016). Furthermore, machine learning approaches have shown considerable potential in mitigating multicollinearity compared to traditional statistical estimators (Chan et al., 2022). For datasets with dependent structures, such as genomic data, novel methods combining grouping and selection have been proposed, outperforming existing variable selection techniques by reducing prediction error while preserving sparsity (Zeng & Xie, 2012).

1.4.5. General effects of multicollinearity on model accuracy

Multicollinearity in regression models can result in increased error rates, reduced predictive accuracy, and unstable coefficient estimates (Efeizomor, 2023; Owoyemi & Bolakale, 2024). Various approaches have been proposed to address this issue, including regularization techniques such as Ridge, LASSO, and Elastic Net regression (Srisa-An, 2021; Schreiber-Gregory et al., 2018). These methods have proven effective in mitigating multicollinearity and enhancing model performance (Efeizomor, 2023; Mahmood, 2024). Machine learning techniques, including RFs and Gradient Boosting, have also demonstrated significant potential in reducing the effects of multicollinearity (Chan et al., 2022). Comparative studies indicate that regularized regression models, particularly Ridge and LASSO, frequently outperform traditional methods such as Ordinary Least Squares (OLS) when multicollinearity is present (Thevaraja & Rahman, 2020; Efeizomor, 2023). However, the choice of method should be guided by the characteristics of the dataset and the specific research objectives (Owoyemi & Bolakale, 2024). Addressing multicollinearity is essential for achieving reliable coefficient estimates and improving the overall interpretability of regression models (Daoud, 2017).

The reviewed studies and their applied computational methods are summarised in Table 1, which provides a comprehensive comparison across classical, soft computing, and hybrid approaches.

Table 1. Comprehensive computational methods for road maintenance cost analysis

	Classical Computation			Soft Computation						
Study	**Multiple Linear Regression**	**Ridge**	**LASSO**	**Artificial Neural Networks**	**Genetic Algorithm**	**Gradient Boosting**	**Support Vector Machines**	**Fuzzy Logic**	**Random Forest**	**Brief Description**
Abdulhafedh, A. (2022)	X	X								Comparison of Ridge Regression and OLS for crash frequency modelling under multicollinearity.
Ahmad, A. (2023)				X						Case study on highway failure and maintenance using ANNs.
Alin, A. (2010)	X									Overview of multicollinearity and its effects in regression.
Aloraini, A. (2017)		X	X							Comparative analysis of Ridge, LASSO, and Elastic Net regularizes for prediction.
Altelbany, S. (2021)		X	X							Evaluation of regularization methods in addressing multicollinearity.
Al-Zwainy, F. M., & Aidan, I. A. A. (2017)				X						ANN-based cost forecasting for highway infrastructure.
Annapoorna, E., et al. (2023)		X	X							Application of Ridge and LASSO in fraud detection using computational methods.
Barua, L., et al. (2021)						X				Gradient Boosting for airport runway pavement deterioration analysis.
Biau, G., & Scornet, E. (2016)									X	Overview of RFs for regression and classification.
Bouabaz, M., & Hamami, M. (2008)				X						ANN-based cost estimation for bridge repair projects.
Chen, J. C., & Wang, X. A. (2017)					X			X		Fuzzy logic and genetic fuzzy models for commodity price prediction.
Dugalam, R., & Prakash, G. (2024)									X	RF algorithm for road health monitoring.
El-Sawalhi, N. I. (2015)							X			SVM-based cost estimation for road projects.
Georgiou, P., et al. (2018)								X		Use of fuzzy logic in pavement roughness prediction.
Fassa, F. (2024)								X		Green infrastructure review focused on fuzzy logic applications.

Study	Classical Computation			Soft Computation						Brief Description
	Multiple Linear Regression	Ridge	LASSO	Artificial Neural Networks	Genetic Algorithm	Gradient Boosting	Support Vector Machines	Fuzzy Logic	Random Forest	
Hapfelmeier, A., & Ulm, K. (2013)									X	Variable selection using RFs techniques.
Makridakis, S., et al. (2018)	X									Comparison of statistical and machine learning forecasting methods.
Guelman, L. (2012)						X				Gradient Boosting for auto insurance cost modelling.
Hamdy, A. (2014)					X			X		Genetic and fuzzy systems for enhancing estimation models.
Costanzi, M., & Cebon, D. (2007)								X		Fuzzy logic for reducing road maintenance costs due to vehicle suspension performance.
Li, Z., & Burgueño, R. (2010)								X		Use of soft computing in bridge evaluation and management.
Raman, R., et al. (2024)									X	IoT-based RF model for real-time road surface assessment.
Rahmawati, F., & Suratman, R. Y. (2022)		X	X							Evaluation of Ridge and LASSO for multicollinear data.
Lu, P., et al. (2020)						X				Gradient Boosting for highway-rail grade crossing crash prediction.
Morris, J. D., & Lieberman, M. G. (2018)	X									Analysis of multicollinearity effects on regression prediction accuracy.
Wang, C., et al. (2022)				X						ANN vs CNN for predicting bridge maintenance costs.
Zhang, Y., et al. (2017)			X							Use of LASSO regression for forecasting highway construction costs.

2. Discussion

This study provides a comprehensive analysis of classical and soft computing methods in predicting road maintenance and repair costs, contributing significantly to the existing body of knowledge. The discussion delves into the comparative performance of these methods, their applicability to diverse datasets, inherent limitations, and sector-specific implications, offering a robust foundation for future research and applications.

Classical computational methods, such as Multiple Linear Regression (MLR), Ridge, and LASSO, have consistently provided a robust framework for predictive modelling. Ridge regression demonstrates notable efficacy in addressing multicollinearity, while LASSO enhances model interpretability by performing variable selection (Aloraini, 2017; Altelbany, 2021). However, these methods are constrained by their reliance on linear assumptions, making them less suitable for datasets characterised by non-linear or intricate relationships (Xi et al., 2023). Additionally, their performance tends to diminish in high-dimensional datasets, where complex interactions between variables are present.

Soft computing techniques have proven indispensable in addressing the limitations of classical methods, particularly in modelling non-linear relationships and managing complex datasets. ANNs have demonstrated superior predictive accuracy in large and intricate datasets, outperforming traditional methods in numerous studies (Bouabaz & Hamami, 2008; Cabrera & Silva, 2023). Similarly, RFs and Gradient Boosting have excelled in handling uncertainties and exploring complex variable interdependencies (Biau & Scornet, 2016; Lu et al., 2020). Furthermore, Fuzzy Logic has shown significant utility in decision-support systems, owing to its ability to manage uncertainty and approximate reasoning (Chen & Wang, 2017).

Multicollinearity presents a critical challenge in regression modelling, adversely affecting prediction accuracy and model stability. While Ridge and LASSO regression offer effective solutions to this issue, soft computing techniques, such as RFs and Gradient Boosting, provide alternative approaches by effectively minimising inter-variable correlations (Chan et al., 2022). Recent work by Annapoorna et al. (2023) demonstrates how ADMM optimisation can enhance the computational efficiency of LASSO and Ridge regression in fraud detection contexts, suggesting potential applications for infrastructure cost modelling. The comparative efficacy of these techniques depends significantly on the structure of the dataset and the specific research objectives.

Environmental factors and material quality play pivotal roles in determining the sustainability and performance of road infrastructure. The incorporation of recycled materials in asphalt mixtures has been shown to enhance cost-efficiency while reducing environmental impacts (Bamigboye et al., 2021). Moreover, climate change, increasing traffic loads, and ageing infrastructure exacerbate the challenges associated with long-term cost estimation and infrastructure management. Addressing these issues necessitates innovative design and construction practices that integrate climate resilience and sustainability principles (Inyim et al., 2016; Kumar & Imam, 2013).

The integration of classical and soft computing techniques within sector-specific applications holds immense potential for developing more comprehensive and effective predictive models. For instance, GAs, when combined with classical methods like Ridge and LASSO, can significantly enhance model accuracy and optimisation (Rehman et al., 2020). Hybrid approaches that leverage the strengths of both computational paradigms are increasingly recognised as a promising avenue for future research and practical application.

Furthermore, beyond the methodological choice, the very nature and categorisation of maintenance activities significantly impact cost prediction models. As detailed in the complementary literature, maintenance operations are not monolithic but are strategically divided into routine, periodic, and emergency interventions (Arshad et al., 2022). Routine maintenance involves regular, low-cost activities to prevent minor defects from worsening, while periodic maintenance encompasses more substantial, planned interventions every few years to preserve structural integrity. Emergency works require immediate response and are funded from a separate contingency budget. This stratification implies that predictive models must be tailored to the specific maintenance type; a model predicting the frequent, low-cost events of routine maintenance will differ fundamentally from one forecasting the sporadic, high-cost needs of periodic rehabilitation or emergency repairs. Ignoring this operational hierarchy can lead to oversimplified and inaccurate cost forecasts. Therefore, future hybrid models should not only integrate computational techniques but also be designed to distinguish between these distinct categories of maintenance expenditure, perhaps by incorporating multi-output architectures or separate sub-models for each maintenance type.

Integration of environmental factors remains underutilised in the reviewed studies. Variables such as climate conditions, material quality, and traffic density, though frequently acknowledged, are rarely fully integrated into predictive models. Fassa's (2024) comprehensive review of green infrastructure highlights emerging applications of fuzzy logic in sustainable road maintenance, particularly for climate adaptation strategies. Future research should explore:

– Incorporating climate-resilient parameters (e.g., freeze–thaw cycles, precipitation patterns) directly into training datasets;
– Using IoT-based real-time monitoring systems to feed environmental data into hybrid predictive models;
– Applying life-cycle assessment frameworks to quantify environmental impacts alongside cost predictions.

The findings of this study highlight several critical gaps in the existing literature. Future research should focus on the following areas:

– **Dataset Diversity:** Analysing datasets from diverse geographical regions and infrastructure types to improve the generalizability of predictive models;

– **Development of Hybrid Models:** Combining classical and soft computing techniques to exploit their complementary strengths and improve predictive accuracy;
– **Integration of Environmental and Socio-Economic Factors:** Incorporating climate change and socio-economic considerations into predictive models to support sustainable infrastructure management.

Despite its comprehensive scope, this study is not without limitations. The diversity of datasets analysed and the contextual applicability of the examined methods may restrict the generalisability of the findings. Nevertheless, the insights presented here lay a solid foundation for the advancement of predictive methodologies in road maintenance and repair cost estimation. By addressing the highlighted gaps, future studies can further refine these approaches, fostering more accurate, efficient, and sustainable infrastructure management solutions.

To synthesise the findings of this narrative review, we have compared the main computational approaches applied in road maintenance cost prediction. As presented in Table 2, each method offers distinct advantages and limitations depending on dataset characteristics and modelling goals. While classical regression techniques (MLR, Ridge, LASSO) remain suitable for smaller and more interpretable datasets, machine learning approaches such as Random Forest, Neural Networks, and Gradient Boosting provide higher predictive performance for large and complex datasets. In contrast, Fuzzy Logic is particularly valuable in contexts where uncertainty and expert judgment play a central role.

Table 2. Comparative summary of methods for road maintenance cost prediction

Method	Best suited for	Advantages	Limitations
MLR	Small, simple datasets	Easy interpretation, widely used	Sensitive to multicollinearity
Ridge / LASSO	Moderate datasets, multicollinearity	Regularization improves stability	Still assumes linear relationships
Random Forest	Large, nonlinear datasets	Handles interactions, robust predictions	Less interpretable
Neural Networks	Very large, complex datasets	High accuracy, nonlinear modelling	Requires more data, less transparent
Fuzzy Logic	Systems with uncertainty	Interpretable, handles vague inputs	Limited generalizability
Gradient Boosting	Medium–large, tabular datasets	High predictive power, flexibility	Risk of overfitting, complex

Conclusions

This study has comprehensively examined the application of classical and soft computing methods in predicting road maintenance and repair costs. By integrating a thorough review of existing literature with critical analyses of methodological advancements, this research has illuminated the strengths, limitations, and sectoral applicability of these techniques. The conclusions drawn from this work provide a foundation for both academic advancement and practical implementation in the field of infrastructure management.

Performance of Classical Methods: Classical methods such as Multiple Linear Regression (MLR), Ridge, and LASSO continue to offer robust solutions for linear relationships, excelling in scenarios where data simplicity and interpretability are prioritised. Ridge regression demonstrates remarkable effectiveness in addressing multicollinearity, while LASSO facilitates variable selection and model sparsity, enhancing interpretability. However, their reliance on linear assumptions limits their applicability in datasets characterised by non-linear relationships or high-dimensionality.

Summary Box 1

Strength: Simple and interpretable.
Ridge: Effective for multicollinearity.
LASSO: Improves model sparsity.
Weakness: Limited in complex datasets.

Advantages of Soft Computing Techniques: Soft computing methods, including ANNs, RFs, Gradient Boosting, Fuzzy Logic, and Genetic Algorithms, have proven indispensable for managing complexity, non-linearity, and uncertainty in data. ANN outperforms classical methods in handling large and intricate datasets, while RFs and Gradient Boosting are particularly effective in variable selection and interdependency management. Fuzzy Logic, with its capacity to model uncertainty, emerges as a critical tool in decision-support systems.

Summary Box 2

Strength: Superior in complex/non-linear systems.
ANNs: High accuracy on complex datasets.
RF & Gradient Boosting: Manage uncertainty and interdependencies.
Fuzzy Logic: Critical for decision support under uncertainty.
More flexible than classical methods.
Weakness: Slow adaptation to dynamic changes.

Hybrid Modelling: Hybrid approaches that integrate classical and soft computing methods demonstrate the potential to exploit the complementary strengths of both paradigms. Techniques combining GAs with Ridge or LASSO exhibit enhanced accuracy and model optimization, suggesting a promising direction for future research and application.

Summary Box 3

Strength: Combines classical and soft computing.
Reported 20–30% accuracy improvement.
Weakness: Long-term validation is limited.
Future work: Validation and IoT integration.

Influence of Environmental and Engineering Factors: The analysis underscores the significant role of environmental and engineering factors in determining the sustainability and performance of road infrastructure. The integration of recycled materials in road construction and the adoption of climate-resilient strategies are essential for reducing costs and environmental impacts, while maintaining infrastructure longevity.

Summary Box 4

Climate, traffic, and material quality is critical.
Recycled materials: Cost and environmental benefits.
Climate-resilient designs: Long-term sustainability.
IoT and LCA integration are future priorities.

Addressing Multicollinearity: Both classical and soft computing methods offer effective solutions to the challenges posed by multicollinearity. While Ridge and LASSO are widely used in classical frameworks, soft computing techniques such as RFs and Gradient Boosting offer alternative strategies for minimising inter-variable correlations and improving model stability.

Summary Box 5

Reduces prediction accuracy and stability.
Ridge & LASSO: Classical solutions.
RF & GB: Alternative strategies.
Method choice depends on dataset structure.

This study makes significant contributions by:

- Comparing classical and soft computing methods for predicting road maintenance costs;
- Identifying critical gaps in the literature and provide actionable recommendations for future research;
- Emphasising the value of hybrid modelling approaches and their potential to enhance accuracy and efficiency in predictive analytics;
- Integrating environmental and socio-economic factors into cost estimation models, thereby supporting sustainable infrastructure management.

The findings of this study have far-reaching implications for policymakers, engineers, and researchers:

- Policymakers can leverage insights from this study to develop evidence-based strategies for infrastructure investment and maintenance planning;
- Engineers can utilise the recommended hybrid modeling techniques to enhance the accuracy and reliability of cost prediction models;
- Researchers are encouraged to explore the integration of environmental and socio-economic dimensions into predictive models, fostering a multidisciplinary approach to infrastructure management.

Future research should address the following key areas:

- **Dataset Diversity:** Conducting analyses on datasets from varied geographic and economic contexts to improve the generalizability of findings;
- **Development of Advanced Hybrid Models:** Exploring innovative combinations of classical and soft computing methods to address complex infrastructure challenges;
- **Sustainability and Resilience:** Investigating the long-term impacts of climate change and material innovations on infrastructure costs and sustainability;
- **Explainable AI in Infrastructure:** Incorporating explainable artificial intelligence (XAI) techniques into predictive models to enhance transparency and decision-making.

While this study offers a comprehensive evaluation of existing methods, it is not without limitations. Limitations of this review include the exclusion of grey literature, the geographical concentration of studies in certain regions, and reliance on secondary data sources. These factors may limit generalisability. However, the diversity of methods reviewed provides valuable guidance for both researchers and practitioners.

The outputs of this narrative review suggest the following tendencies regarding the application of different computational approaches in road maintenance cost estimation and forecasting. Classical methods such as MLR, Ridge, and LASSO remain appropriate when datasets are small, and model interpretability is prioritised. In contrast, Random Forest, Neural Networks, and Gradient Boosting offer higher predictive accuracy for large and complex datasets, though at the expense of interpretability. Fuzzy Logic may be most useful where qualitative uncertainty and expert knowledge are involved.

This study provides a robust framework for advancing predictive methodologies in road maintenance and repair cost estimation. By addressing the limitations and exploring the identified opportunities, future research can enhance the accuracy, efficiency, and sustainability of infrastructure management practices. Through the adoption of hybrid modelling approaches and the integration of environmental and socio-economic factors, this field has the potential to make a substantial impact on global infrastructure sustainability.

Acknowledgements

The corresponding author deeply appreciates Seda Arslan and Ali Asker Yalçın for their constant support and motivation. The authors sincerely express their gratitude to esteemed advisors and committed colleagues in the İskenderun Municipality Parks and Gardens Department, whose essential guidance and support were crucial for the successful completion of this project.

This study is a part of the Doctoral Thesis titled “Determination of Highway Maintenance and Repair Costs with Soft Computing and Classical Techniques” of Haydar Gundogdu, which is being prepared at the Institute of Graduate Studies, Civil Engineering Program, at Iskenderun Technical University.

Disclosure Statement

The authors hereby declare that there are no financial, professional, or personal conflicts of interest that could have influenced the design, execution, or presentation of the work detailed in this article.

REFERENCES

Abdulhafedh, A. (2022). Modeling vehicle crash frequency when multicollinearity exists in vehicle crash data: ridge regression versus ordinary least squares linear regression. *Open Access Library Journal*, *9*(6), 1–17.

Abed, Y. G., Hasan, T. M., & Zehawi, R. N. (2022). Cost prediction for roads construction using machine learning models. *International Journal of Electrical and Computer Engineering Systems*, *13*(10), 927–936. https://doi.org/10.32985/ijeces.13.10.8

Adefabi, A., Olisah, S., Obunadike, C., Oyetubo, O., Taiwo, E., & Tella, E. (2023). Predicting accident severity: an analysis of factors affecting accident severity using random forest model. *arXiv preprint* 2310.05840. https://doi.org/10.48550/arXiv.2310.05840

Ahmad, A. (2023). Analyzing the highway failure and their maintenance: A case study. *International Journal of Scientific Research in Engineering and Management*, *7*. https://doi.org/10.55041/IJSREM24615

Alin, A. (2010). Multicollinearity. *WIREs Computational Statistics*, *2*(3), 370–374. https://doi.org/10.1002/wics.84

AlKheder, S. (2022). Environment-friendly recycled asphalt pavement design for road maintenance applications. *Environment, Development and Sustainability*, *24*, 14210–14234. https://doi.org/10.1007/s10668-021-02027-2

Aloraini, A. (2017). On the prediction accuracies of three most known regularizers: Ridge regression, the LASSO estimate, and elastic net regularization methods. *International Journal of Artificial Intelligence and Applications*, *8*(6), 29–36. https://doi.org/10.5121/IJAIA.2017.8603

Altelbany, S. (2021). Evaluation of ridge, elastic net and lasso regression methods in precedence of multicollinearity problem: a simulation study. *Journal of Applied Economics and Business Studies*, *5*(1), 131–142. https://doi.org/10.34260/jaebs.517

Al-Zwainy, F. M., & Aidan, I. A. A. (2017). Forecasting the cost of structure of infrastructure projects utilizing artificial neural network model (highway projects as case study). *Indian Journal of Science and Technology*, *10*(20), 1–12.

Annapoorna, E., Sujil, S. V., & Gopakumar, G. (2023, November). ADMM: Computational Analysis analysis of Lasso and Ridge on Fraud fraud Detectiondetection. *2023 2nd International Conference on Futuristic Technologies (INCOFT)*, Belagavi, Karnataka, India, (pp. 1–6). https://doi.org/10.1109/INCOFT60753.2023.10425471

Arshad, A. K., Shaffie, E., Kamal, M. I. A., Hussain, M. Z., Rais, N. M. (2022). Application of pavement evaluation for road maintenance and rehabilitation. R. Hassan, R., et al. (Eds.), *Green Infrastructure* (pp. 159–177). Springer. https://doi.org/10.1007/978-981-16-6383-3_10

Atapattu, C. N., Domingo, N. D., & Sutrisna, M. (2022, November). Statistical cost modelling for preliminary stage cost estimation of infrastructure projects. *IOP Conference Series Earth and Environmental Science*, *1101*(5), 052031. https://doi.org/10.1088/1755-1315/1101/5/052031

Awwad, M. (2021). Studying the effects of roads geometry and design parameters on the pavement drainage system. *Civil Engineering Journal*, *7*(1), 49–58. https://doi.org/10.28991/cej-2021-03091636

Bamigboye, G. O., Bassey, D. E., Olukanni, D. O., Ngene, B. U., Adegoke, D., Odetoyan, A. O., Kareem, M. A., ...Enabulele, D. O., & Nworgu, A. T. (2021). Waste materials in highway applications: An overview on generation and utilization implications on sustainability. *Journal of Cleaner Production*, *283*, 124581. https://doi.org/10.1016/j.jclepro.2020.124581

Barakat, S. (2023). Review of classical methods and variables selection in case of multicollinearity: A case study with real-data. *Trade and financing*, *4*(43), 225–246. https://doi.org/10.21608/caf.2023.336191

Barua, L., Zou, B., Noruzoliaee, M., & Derrible, S. (2021). A gradient boosting approach to understanding airport runway and taxiway pavement deterioration. *International Journal of Pavement Engineering*, *22*(13), 1673–1687. https://doi.org/10.1080/10298436.2020.1714616

Biau, G., & Scornet, E. (2016). A random forest guided tour. *Test*, *25*, 197–227. https://doi.org/10.1007/s11749-016-0481-7

Bouabaz, M., & Hamami, M. (2008). A cost estimation model for repair bridges based on artificial neural network. *American Journal of Applied Sciences*, *5*(4), 334–339. https://doi.org/10.3844/ajassp.2008.334.339

Cabrera, M. S., & Silva, D. L. (2023, June). Neural network-based prediction of the performance budget for road defects and highway maintenance. *2023 IEEE 3rd International Conference on Software Engineering and Artificial Intelligence (SEAI)*, Xiamen, China, (pp. 63–67). https://doi.org/10.1109/SEAI59139.2023.10217472

Chan, J. Y. L., Leow, S. M. H., Bea, K. T., Cheng, W. K., Phoong, S. W., Hong, Z. W., & Chen, Y. L. (2022). Mitigating the multicollinearity problem and its machine learning approach: a review. *Mathematics*, *10*(8), 1283. https://doi.org/10.3390/math10081283

Chen, J. C., & Wang, X. A. (2017, December). Development of fuzzy logic and genetic fuzzy commodity price prediction systems – An industrial case study. *2017 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM)*, Singapore, (pp. 432–435). https://doi.org/10.1109/IEEM.2017.8289927

Costanzi, M., & Cebon, D. (2007). An investigation of the effects of lorry suspension performance on road maintenance costs. *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science*, *221*(11), 1265–1277. https://doi.org/10.1243/09544062JMES639

Daoud, J. I. (2017, December). Multicollinearity and regression analysis. *Journal of Physics: Conference Series*, *949*(1), 012009. https://doi.org/10.1088/1742-6596/949/1/012009

de Sá Ferreira, F. (2019). The impact of fuzzy logic on logistics optimization: A comprehensive review. *Ciências Sociais Aplicadas*, *23*. https://doi.org/10.69849/revistaft/ar10201901011035

Dugalam, R., & Prakash, G. (2024). Development of a random forest based algorithm for road health monitoring. *Expert Systems with Applications*, *251*, Article 123940. https://doi.org/10.1016/j.eswa.2024.123940

Efeizomor, R. O. (2023). A comparative study of methods of remedying multicolinearity. *American Journal of Theoretical and Applied Statistics*, *14*(1), 87–91. https://doi.org/10.11648/j.ajtas.20231204.14

El-Sawalhi, N. I. (2015). Support vector machine cost estimation model for road projects. *Journal of Civil Engineering and Architecture*, *9*(9), 1115–1125. https://www.davidpublisher.com/Public/uploads/Contribute/561c6f5db2dff.pdf

Etemadi, S., & Khashei, M. (2021). Etemadi multiple linear regression. *Measurement*, *186*, 110080. https://doi.org/10.1016/j.measurement.2021.110080

Fassa, F. (2024, May). Green infrastructure issues 2016–2021: Literature review. *AIP Conference Proceedings*, *2891*(1), 050014. https://doi.org/10.1063/5.0200958

Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. *Annals of Statistics*, *29*(5), 1189–1232. https://doi.org/10.1214/aos/1013203451

Garg, A., & Tai, K. (2013). Comparison of statistical and machine learning methods in modelling of data with multicollinearity. *International Journal of Modelling, Identification and Control*, *18*(4), 295–312. https://doi.org/10.1504/IJMIC.2013.053535

Georgiou, P., Plati, C., & Loizos, A. (2018). Soft computing models to predict pavement roughness: A comparative study. *Advances in Civil Engineering*, *2018*(1), 5939806. https://doi.org/10.1155/2018/5939806

Gibby, R., Kitamura, R., & Zhao, H. (1990). Evaluation of truck impacts on pavement maintenance cost. *Transportation Research Record*, *1262*, 48–56. https://onlinepubs.trb.org/Onlinepubs/trr/1990/1262/1262-006.pdf

Goldberg, D. E. (2013). *Genetic algorithms*. Pearson education India.

Grefenstette, J. J. (Ed). (2014). *Proceedings of the First International Conference on Genetic Algorithms and their Applications*. Psychology Press.

Guelman, L. (2012). Gradient boosting trees for auto insurance loss cost modeling and prediction. *Expert Systems with Applications*, *39*(3), 3659–3667. https://doi.org/10.1016/j.eswa.2011.09.058

Hamdy, A. (2014). Genetic fuzzy system for enhancing software estimation models. *International Journal of Modeling and Optimization*, *4*(3), 227–232. https://doi.org/10.7763/IJMO.2014.V4.378

Hapfelmeier, A., & Ulm, K. (2013). A new variable selection approach using Random Forests. *Computational Statistics & Data Analysis*, *60*, 50–69. https://doi.org/10.1016/j.csda.2012.09.020

Hashem, H. A. (2022). A comparative study of some variables selection methods in high dimensional multiple liner regression via simulation. *Journal of Duhok University*, *25*(2), 601–613. https://doi.org/10.26682/sjuod.2022.25.2.53

Herawati, N. N., Wijayanti, N. A., Sutrisno, N. A., Nusyirwan, N., & Misgiyati, N. (2024). The performance of Ridge Regression, LASSO, and Elastic-Net in controlling multicollinearity: a simulation and application. *Journal of Modern Applied Statistical Methods*, *23*(2). https://doi.org/10.56801/jmasm.v23.i2.4

Hou, Y., Edara, P., & Sun, C. (2012, September). A genetic fuzzy system for modeling mandatory lane changing. *2012 15th International IEEE Conference on Intelligent Transportation Systems* (pp. 1044–1048). https://doi.org/10.1109/ITSC.2012.6338877

Hu, Y. H., Yu, S. C., Qi, X., Zheng, W. J., Wang, Q. Q., & Yao, H. Y. (2019). An overview of multiple linear regression model and its application. *Zhonghua yu Fang yi xue za zhi* [*Chinese Journal of Preventive Medicine*], *53*(6), 653–656.

Inyim, P., Pereyra, J., Bienvenu, M., & Mostafavi, A. (2016). Environmental assessment of pavement infrastructure: A systematic review. *Journal of Environmental Management*, *176*, 128–138. https://doi.org/10.1016/j.jenvman.2016.03.042

Kilinc, B., Aşikgil, B., Erar, A., & Yazici, B. (2016). Variable selection with genetic algorithm and multivariate adaptive regression splines in the presence of multicollinearity. *International Journal of Advanced and Applied Sciences*, *3*(12), 26–31. https://doi.org/10.21833/ijaas.2016.12.004

Koukol, M., Zajíčková, L., Marek, L., & Tuček, P. (2015). Fuzzy logic in traffic engineering: a review on signal control. *Mathematical Problems in Engineering*, *2015*(1), Article 979160. https://doi.org/10.1155/2015/979160

Kumar, P., & Imam, B. (2013). Footprints of air pollution and changing environment on the sustainability of built infrastructure. *Science of the Total Environment*, *444*, 85–101. https://doi.org/10.1016/j.scitotenv.2012.11.056

Lagari, P. L., Tsoukalas, L. H., Safarkhani, S., & Lagaris, I. E. (2021). Eliminating multicollinearity issues in neural network ensembles: Incremental, negatively correlated, optimal convex blending. *arXiv preprint 2104.14715*. https://doi.org/10.48550/arXiv.2104.14715

Li, J., Yin, G., Wang, X., & Yan, W. (2022). Automated decision making in highway pavement preventive maintenance based on deep learning. *Automation in Construction*, *135*, 104111. https://doi.org/10.1016/j.autcon.2021.104111

Li, Z., & Burgueño, R. (2010). Using soft computing to analyze inspection results for bridge evaluation and management. *Journal of Bridge Engineering*, *15*(4), 430–438. https://doi.org/10.1061/(ASCE)BE.1943-5592.0000072

Li, Z., Lang, L., Sun, G., Cai, Z., & Luo, Z. (2023, December). Enhancing multiple linear regression for price prediction: A PCA-integrated approach. *2023 4th International Conference on Computer, Big Data and Artificial Intelligence (ICCBD+ AI)*, Guiyang, China, (pp. 337–341). https://doi.org/10.1109/ICCBD-AI62252.2023.00063

Li, Z., Zheng, Z., & Washington, S. (2019). Short-term traffic flow forecasting: A component-wise gradient boosting approach with hierarchical reconciliation. *IEEE Transactions on Intelligent Transportation Systems*, *21*(12), 5060–5072. https://doi.org/10.1109/TITS.2019.2948381

Liljenström, C., Björklund, A., & Toller, S. (2022). Including maintenance in life cycle assessment of road and rail infrastructure— a literature review. *The International Journal of Life Cycle Assessment*, *27*(2), 316–341. https://doi.org/10.1007/s11367-021-02012-x

Lu, P., Zheng, Z., Ren, Y., Zhou, X., Keramati, A., Tolliver, D., & Huang, Y. (2020). A gradient boosting crash prediction approach for highway-rail grade crossing crash analysis. *Journal of Advanced Transportation*, *2020*(1), 6751728. https://doi.org/10.1155/2020/6751728

Mahmood, S. H. (2024). Estimating models and evaluating their efficiency under multicollinearity in multiple linear regression: A comparative study. *Zanco Journal of Human Sciences*, *28*(5), 264–277. https://doi.org/10.21271/zjhs.28.5.17

Makridakis, S., Spiliotis, E., & Assimakopoulos, V. (2018). Statistical and machine learning forecasting methods: Concerns and ways forward. *PlOS One*, *13*(3), e0194889. https://doi.org/10.1371/journal.pone.0194889

Manogaran, G., Rodrigues, J. J., Kozlov, S. A., & Manokaran, K. (20212022). Conditional support-vector-machine-based shared adaptive computing model for smart city traffic management. *IEEE Transactions on Computational Social Systems*, *9*(1), 174–183. https://doi.org/10.1109/TCSS.2021.3051330

Menguc, K., Aydin, N., & Yilmaz, A. (2023). A data driven approach to forecasting traffic speed classes using extreme gradient boosting algorithm and graph theory. *Physica A: Statistical Mechanics and its Applications*, *620*, 128738. https://doi.org/10.1016/j.physa.2023.128738

Morris, J. D., & Lieberman, M. G. (2018). Multicollinearity's effect on regression prediction accuracy with real data structures. *General Linear Model Journal*, *44*(1), 29–34. https://doi.org/10.31523/glmj.044001.004

Nabil, M., Aly, M., & Atiya, A. (2014). LABR: A large scale arabic sentiment analysis benchmark. *arXiv preprint* 1411.6718. https://doi.org/10.48550/arXiv.1411.6718.

Narayanan, R. M., Annapoorna, E., Sujil, S. V., & Gopakumar, G. (2023, November). ADMM: Computational analysis of Lasso and Ridge on fraud detection. *2023 2nd International Conference on Futuristic Technologies (INCOFT)*, Belagavi, Karnataka, India, (pp. 1–6). https://doi.org/10.1109/INCOFT60753.2023.10425471

Naveed, S. A., Farooq, U., Raza, M. A., Rehman, Z. U., Saleem, M., & Ghazal, T. M. (2024). Enhancing traffic flow and congestion management in smart cities utilizing SVM-based linear regression approach. *International Journal of Advanced and Applied Sciences*, *11*(10), 166–175. https://doi.org/10.21833/ijaas.2024.10.019

Nyame, M. (2024). Impact of geotechnical engineering on infrastructure lifespan and maintenance costs. *Journal of Scientific Research and Reports*, *30*(9), 217–226. https://doi.org/10.9734/jsrr/2024/v30i92346

Obite, C. P., Olewuezi, N. P., Ugwuanyim, G. U., & Bartholomew, D. C. (2020). Multicollinearity effect in regression analysis: A feed forward artificial neural network approach. *Asian Journal of Probability and Statistics*, *6*(1), 22–33. https://doi.org/10.9734/ajpas/2020/v6i130151

Omer, P. A. (2022). Improving prediction accuracy of Lasso and Ridge regression as an alternative to LS regression to identify variable selection problems. https://www.researchgate.net/publication/385137550_Improving_Prediction_Accuracy_of_Lasso_and_Ridge_Regression_as_an_Alternative_to_LS_Regression_to_Identify_Variable_Selection_Problems

Owoyemi, Q. A., & Bolakale, A. (2024). Comparative analysis of some linear predictive models in the presence of multicollinearity. *International Journal of Advanced Statistics and Probability*, *11*(1). https://doi.org/10.14419/r2bqgv16

Oyeyemi, G. M., Ogunjobi, E. O., & Folorunsho, A. I. (2015). On performance of shrinkage methods – a A Monte Carlo Studystudy. *International Journal of Statistics and Applications*, *5*(2), 72–76. https://www.researchgate.net/publication/313814555_On_Performance_of_Shrinkage_Methods_-_A_Monte_Carlo_Study

Pais, J. C., Figueiras, H., Pereira, P., & Kaloush, K. (2019). The pavements cost due to traffic overloads. *International Journal of Pavement Engineering*, *20*(12), 1463–1473. https://doi.org/10.1080/10298436.2018.1435876

Paul, R. K. (2006). Multicollinearity: Causes, effects and remedies.

Petrusheva, S., Car-Pušić, D., & Zileska-Pancovska, V. (2019). Support vector machine based hybrid model for prediction of road structures construction costs. *IOP Conference Series Earth and Environmental Science*, *222*(1), 012010. https://doi.org/10.1088/1755-1315/222/1/012010

Portnoy, S., & Koenker, R. (1997). The Gaussian hare and the Laplacian tortoise: computability of squared-error versus absolute-error estimators. *Statistical Science*, *12*(4), 279–300. https://doi.org/10.1214/ss/1030037960

Radhika, K., Valarmathy, S., Selvarasu, S., Bashkaran, K., & Srinivasan, C. (2023, November). Predictive Road Sign Maintenance using Random Forest Regression and IoT Data. *2023 International Conference on Sustainable Communication Networks and Application (ICSCNA)* (pp. 348–353). https://doi.org/10.1109/ICSCNA58489.2023.10370728

Rahmawati, F., & Suratman, R. Y. (2022). Performa Regresi Ridge dan Regresi Lasso pada Data dengan Multikolinearitas. *Leibniz: Jurnal Matematika*, *2*(2), 1–10. https://doi.org/10.59632/leibniz.v2i2.176

Raman, R., Vekariya, V., Gurpur, S., Narayana, K. E., & Murugan, S. (2024, November). Real-time road surface assessment through IoT data analysis and random forest. *2024 8th International Conference on Electronics, Communication and Aerospace Technology (ICECA)*, Coimbatore, India, (pp. 290–295). https://doi.org/10.1109/ICECA63461.2024.10800879 IEEE.

Ravishankar, S., & Latha, P. (2012). Software cost estimation using fuzzy logic. *International Conference on Recent Trends in Computational Methods, Communication and Controls* (ICON3C). https://scispace.com/pdf/software-cost-estimation-using-fuzzy-logic-3d786gzwf8.pdf

Rehman, G., Zhang, G., Rahman, M.U., Rahman, N.U., Usman, T. and Imraz, M. (2020) The Engineering Assessments and Potential Aggregate Analysis of Mesozoic Carbonates of Kohat Hills Range, KP, Pakistan. *Acta Geodaetica et Geophysica*, *55*, 477–493.

Reznychenko, T., Uglickich, E., & Nagy, I. (2024, May). Accuracy comparison of logistic regression, random forest, and neural networks applied to real MaaS data. *2024 Smart City Symposium Prague (SCSP)*, Prague, Czech Republic, (pp. 1–5). https://doi.org/10.1109/SCSP61506.2024.10552715

Rohani, A., Abbaspour-Fard, M. H., & Abdolahpour, S. (2011). Prediction of tractor repair and maintenance costs using Artificial Neural Network. *Expert Systems with Applications*, *38*(7), 8999–9007. https://doi.org/10.1016/j.eswa.2011.01.118

Ryu, N., Kim, H., & Kang, P. (2016). Evaluating variable selection techniques for multivariate linear regression. *Journal of Korean Institute of Industrial Engineers*, *42*(5), 314–326. https://doi.org/10.7232/JKIIE.2016.42.5.314

Sağlık, A., & Güngör, G. (2008). *Flexible pavement design guide for highways*, (Topic #625-010-002). General Directorate of Highways, Department of Technical Research, Pavement Branch, Ankara, Turkey. https://fdotwww.blob.core.windows.net/sitefinity/docs/default-source/roadway/pm/publications/fpdm201801.pdf?sfvrsn=149199e9_2

Salih, A. M. (2024). Explainable artificial intelligence and multicollinearity: A Mini mini review of current approaches. *arXiv preprint 2406.11524*. https://doi.org/10.48550/arXiv.2406.11524

Salman, H. A., Kalakech, A., & Steiti, A. (2024). Random Forest algorithm overview. *Babylonian Journal of Machine Learning*, *2024*, 69–79. https://doi.org/10.58496/BJML/2024/007

Sandri, M., & Zuccolotto, P. (2006, August). Variable selection using random forests. In S. Zani, A. Cerioli, M. Riani, & M. Vichi, (Eds.) *Data Analysis, Classification and the Forward Search. Studies in Classification, Data Analysis, and Knowledge Organization*. Springer, Berlin, Heidelberg (pp. 263–270). https://doi.org/10.1007/3-540-35978-8_30

Santos, J., Ferreira, A., & Flintsch, G. (2019). An adaptive hybrid genetic algorithm for pavement management. *International Journal of Pavement Engineering*, *20*(3), 266–286. https://doi.org/10.1080/10298436.2017.1293260

Santos, G. S., Marotta, L. I. M., Clemente, G. A. B. F., Lima, G. de M., Scarano, D., & Oliveira, C. R. de. (2024). Concrete roads: Sustainable paving. *Environmental & Social Management Journal*, *Revista de Gestão - RGSA, 18*(10), e08596. https://doi.org/10.24857/rgsa.v18n10-095

Sapkota, K., Yaghoubi, E., Wasantha, P. L. P., Van Staden, R., & Fragomeni, S. (2023). Mechanical characteristics and durability of HMA made of recycled aggregates. *Sustainability*, *15*(6), 5594. https://doi.org/10.3390/su15065594

Schreiber-Gregory, D. N. (2018). Ridge Regression and multicollinearity: An in-depth review. *Model Assisted Statistics and Applications*, *13*(4), 359–365. https://doi.org/10.3233/MAS-180446

Smith, A. S., Odolinski, K., Hossein-Nia, S., Jönsson, P. A., Stichel, S., Iwnicki, S., & Wheat, P. (2021). Estimating the marginal maintenance cost of different vehicle types on rail infrastructure. *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, *235*(10), 1191–1202. https://doi.org/10.1177/0954409721991309

Soni, J., & Gujar, R. (2024). Classifying flexible pavement defects using hybrid machine learning approach. *International Journal of Electrical and Computer Engineering (IJECE)*, *34*(2), 1072–1080. http://doi.org/10.11591/ijeecs.v34.i2.pp1072-1080

Srisa-An, C. (2021, August). Guideline of collinearity-avoidable regression models on time-series analysis. *2021 2nd International Conference on Big Data Analytics and Practices (IBDAP)*, Bangkok, Thailand, (pp. 28–32). https://doi.org/10.1109/IBDAP52511.2021.9552165

Thevaraja, M., & Rahman, A. (2020). Assessing robustness of regularized regression models with applications. In Xu, J., Ahmed, S., Cooke, F., Duca, G. (eds) *Proceedings of the Thirteenth International Conference on Management Science and Engineering Management. ICMSEM 2019. Advances in Intelligent Systems and Computing*, *1001*. https://doi.org/10.1007/978-3-030-21248-3_30

Thorat, M., & Birajdar, B. V. (2015). Cost overrun assessment model in highway construction projects using fuzzy uncertainty analysis. *International Journal of Engineering Research & Technology (IJERT)*, *6*(4), 568–574. https://www.ijert.org/cost-overrun-assessment-model-in-highway-construction-projects-using-fuzzy-uncertainty-analysis

Upendra, S., Abbaiah, R., & Balasiddamuni, P. (2023). Multicollinearity in multiple linear regression: Detection, consequences, and remedies. *International Journal for Research in Applied Science and Engineering Technology*, *11*(9), 1047–1061. https://doi.org/10.22214/ijraset.2023.55786

Vardon, P. J. (2015). Climatic influence on geotechnical infrastructure: a review. *Environmental Geotechnics*, *2*(3), 166–174. https://doi.org/10.1680/envgeo.13.00055

Wang, C., Yao, C., Zhao, S., Zhao, S., & Li, Y. (2022). A comparative study of a fully- connected artificial neural network and a convolutional neural network in predicting bridge maintenance costs. *Applied Sciences*, *12*(7), 3595. https://doi.org/10.3390/app12073595

Woldemariam, W., Murillo-Hoyos, J., & Labi, S. (2016). Estimating annual maintenance expenditures for infrastructure: Artificial neural network approach. *Journal of Infrastructure Systems*, *22*(2), 04015025. https://doi.org/10.1061/(ASCE)IS.1943-555X.0000280

Xi, L. J., Guo, Z. Y., Yang, X. K., & Ping, Z. G. (2023). Application of LASSO and its extended method in variable selection of regression analysis. *Zhonghua yu fang yi xue za zhi* [*Chinese Journal of Preventive Medicine*], *57*(1), 107–111.

Yıldırım, H. (2024). The multicollinearity effect on the performance of machine learning algorithms: Case examples in healthcare modelling. *Academic Platform Journal of Engineering and Smart Systems*, *12*(3), 68-–80. https://doi.org/10.21541/apjess.1371070

Ying, J., Zhang, X., Zhang, Y., & Bilan, S. (2022). Green infrastructure: Systematic literature review. *Economic research-Ekonomska istraživanja*, *35*(1), 343–366. https://doi.org/10.1080/1331677X.2021.1893202

Yu, B., Chen, Q., Li, N., Wang, Y., Li, L., Cai, M., … Zhang, W., Gu, T., Zhu, R., Zeng, H., Zhang, Y., & Jiang, J. (2024). Life cycle assessment of urban road networks: Quantifying carbon footprints and forecasting future material stocks. *Construction and Building Materials*, *428*, 136280. https://doi.org/10.1016/j.conbuildmat.2024.136280

Zeng, L., & Xie, J. (2012). Group variable selection for data with dependent structures. *Journal of Statistical Computation and Simulation*, *82*(1), 95–106. https://doi.org/10.1080/00949655.2010.529812

Zhang, Y., Cai, Y., & Wu, H. (2015, April). Research on cost estimation of highway project based on the GA-BP algorithm. *Proceedings of the 19th International Symposium on Advancement of Construction Management and Real Estate* (pp. 451–462). Berlin, Heidelberg: Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-662-46994-1_37

Zhang, Y., Minchin Jr, R. E., & Agdas, D. (2017). Forecasting completed cost of highway construction projects using LASSO regularized regression. *Journal of Construction Engineering and Management*, *143*(10), 04017071. https://doi.org/10.1061/(ASCE)CO.1943-7862.0001378

Zhou, X., Lu, P., Zheng, Z., Tolliver, D., & Keramati, A. (2020). Accident prediction accuracy assessment for highway-rail grade crossings using random forest algorithm compared with decision tree. *Reliability Engineering & System Safety*, *200*, 106931. https://doi.org/10.1016/j.ress.2020.106931